Sağlık Araştırmalarında Kullanılan Temel İstatistik Yöntemleri ve Hipotez Testleri

Yrd. Doç. Dr . Emre ATILGAN

BİYOİSTATİSTİK

- İstatistiğin biyoloji, tıp ve diğer sağlık bilimlerinde kullanımı "**biyoistatistik**" olarak tanımlanmaktadır.
- Biyoistatistiğin sağlık bilimlerindeki kullanım alanları ise genel olarak hizmet planlaması, tanı ve tedavi işlemleri, toplumsal değişimlerin incelenmesi, koruyucu hizmetler, biyolojik, morfolojik ve fizyolojik özelliklerin tanımlanması, bilimsel çalışmalar ve hizmetin Ölçümlenmesi şeklinde sınışanmaktadır.

İstatistiksel Yöntemler	Tanımlatıcı İstatistikleri	Frekans Dağılım Tablosu	
		Merkezi Eğilim Ölçüleri	Aritmetik ortalama Medyan (Ortanca) Mod (Tepe değeri)
		Merkezi Dağılım Ölçüleri	Varyans ve Standart Sapma Değişim Aralığı Değişim Katsayısı
	Yorumlayıcı İstatistikler	Tahminleme	Nokta Tahmini Aralık Tahmini
		Hipotez Testleri	1.Sıfır Hipotezi (H_0) 2. Alternatif Hipotez (H_1) 3. Hata Düzeyi (α) 4. Test ve Test İstatistiği

İstatistiksel Testler

Parametrik Testler

Gözlemlerin bağımsız olması (Birinin seçimi diğerini etkilememeli)

Gözlemlerin normal dağılım(çan eğrisi) gösteren kitlelerden seçilmesi

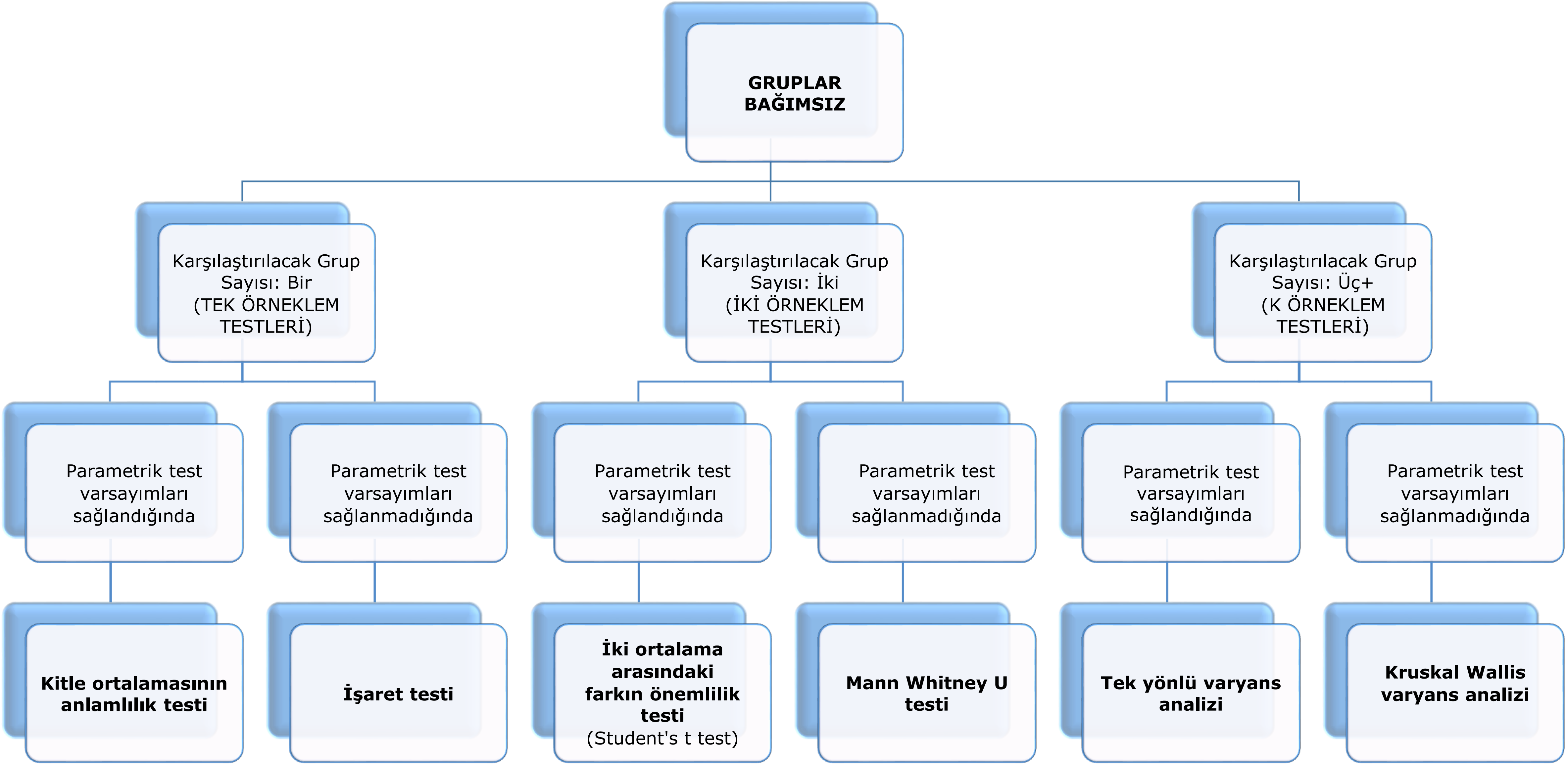
Örneklem hacminin en az 30 olması (n ≥30) gereklidir.

Parametrik Olmayan Testler

Parametrik test varsayımlarının sağlanamadığı durumlarda, diğer bir ifade ile sayısal ölçüm değerleri yerine skorlamalar ve benzeri sıralamalar kullanıldığında, ve örneklem hacmi 10'dan az olduğunda (n≤ 10) parametrik olmayan istatistiksel teknikler yardımıyla analizler tamamlanır.

Parametrik olmayan testler daha az duyarlı ölçme düzeylerinde kullanılabildiklerinden ve daha az varsayım gerektirdiklerinden uygulama alanları parametrik testlere göre daha geniştir

Veri Tipi: Sayısal (Ölçümle Elde Edilmiş)



İşaret Testi

- ilgilenilen değişkenin dağılımının medyan değerinin belirli bir değere eşit olup olmadığının testinde kullanılır.
- İşaret testi;
  - Tek örneklem t testi yerine,
  - Eşlenmiş t testi yerine,
  - Sayısal bir ölçeklemenin mümkün olmadığı ancak gözlemlerin bir şekilde sıralanabildiği kategorik verilerde kullanılabilir.
- Hipotezler:
  - $H_0: M = a$
  - $H_1: M \neq a$

- $\dot{I}_H > \dot{I}_{tablo}$ → H_0 red.

Mann-Whitney U Testi

- Veri parametrik test varsayımlarını sağlamıyor ise; İki Ortalama Arasındaki Farkın Önemlilik Testi yerine kullanılabilecek en güçlü testtir.

- **H_0 hipotezi:**

"İki ortalama arasında fark yoktur" şeklinde değil,

"İki dağılım arasında fark yoktur" şeklinde kurulur.

- **Her iki gruptaki denek sayıları 20 ya da daha az ise:**

$U_H > U_{tablo} \rightarrow H_0$ red.

- **Her iki gruptaki denek sayıları 20 ya da daha az ise:**

$z_H > z_{tablo} \rightarrow H_0$ red.

Tek Yönlü Varyans Analizi

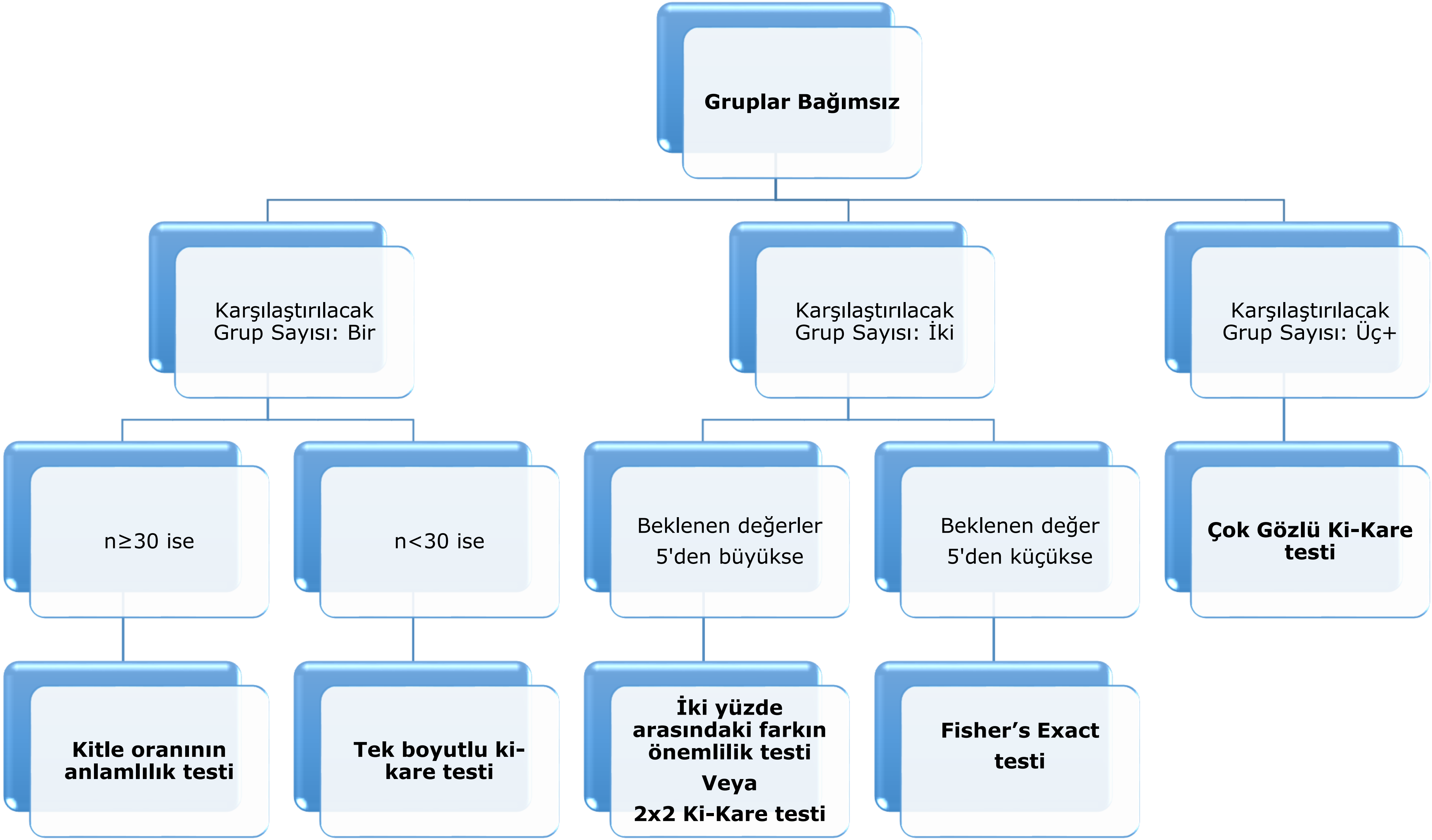
- Parametrik test varsayımları sağlandığında, ölçümle belirtilen bir değişken yönünden **ikiden fazla bağımsız grubun** ortalamaları arasında fark olup olmadığını test etmek için kullanılır. İki ortalama arasındaki farkın anlamlılık testi için gerekli varsayımlar varyans analizi için de geçerlidir.
- Varsayımlar:
  - Karşılaştırılacak gruplar normal dağılım göstermeli
  - Grupların varyansları homojen olmalı
  - Gruplar birbirinden bağımsız olmalı
- Hipotezler:
  - H_0: $\mu_1 = \mu_2 = \mu_3 = ... = \mu_k$
  - H_1: En az bir μ_i farklıdır
- $F_{hesaplanan} > F_{\alpha;(n-k);(n-1)}$ → H0 red. (n: toplam gözlem sayısı, k:grup sayısı)

Kruskall - Wallis Testi

- Tek yönlü varyans analizinin parametrik olmayan karşılığıdır. Veriler ölçümle belirtildiği halde parametrik test varsayımları sağlanmıyorsa (gözlem sayısı az ya da gruplar normal dağılmıyor ise) Kruskal-Wallis testi kullanılır.
- Hipotezler:
  - H_0: İlgili değişken açısından gruplar arasında fark yoktur.
  - H_1: İlgili değişken en az bir grupta farklıdır.
- Testin aşamaları şu şekilde gerçekleşir:

1. k grubun n_1, n_2,…, n_k gözlemleri tek bir değişken altında küçükten büyüğe sıralanır. Tüm gözlemlere sıra numarası verilir.
2. k grubun sıra numaraları ayrı ayrı toplanır
3. Test istatistiği hesaplanır.
4. $KW_{hesaplanan} > \chi^2_{(k-1)}$ → H_0 red.

Değişken Tipi: Kategorik (Sayımla Elde Edilmiş)



Tek Boyutlu Ki – Kare Testi (Ki-Kare İyi Uyum Testi)

- İlgilenilen değişken sınıflayıcı ölçek ile ölçümlenmiş, iki ya da daha fazla kategori söz konusu ise bu değişkenin aldığı değerlerin teorik olarak beklenen değerler ile olan uyumunu test etmek için kullanılır.
- Bu yöntem, örneklemdeki denek sayısı $n < 30$ olduğunda kitle oranının anlamlılık testi yerine kullanılır. Bu yöntemde test istatistiği olarak ki-kare kullanılır.
- Hipotezler:

  H_0: İlgilenen özellik var

  H_1: İlgilenen özellik yok

$\chi^2_{(H)} > \chi^2_{(m-1)} \rightarrow H_0$ red (m: Kategori sayısı)

Tek Boyutlu Ki – Kare Testi (Ki-Kare İyi Uyum Testi)

Örnek: Çocuk felci aşılama programında bir bölgedeki aşılama oranının 0.80 olduğu düşünülmektedir. Bu bölgeden rasgele seçilen 25 çocuktan 18'inin aşılanmış olduğu saptandığına göre bölgedeki aşılama oranının 0.80 olduğu söylenebilir mi?

	G	B
Aşılanmış	18	20
Aşılanmamış	7	5
Toplam	25	25

$$\chi^2 = \frac{(18-20)^2}{20} + \frac{(7-5)^2}{5} = 1$$

$$\alpha = 0.05 \text{ için} \qquad \chi^2_{1,0.05} = 3.841$$

$\chi^2_{(H)} < \chi^2_{(m-1)} \rightarrow H_0$ kabul

Bölgedeki aşılama oranının 0.80 olduğu söylenebilir.

Tek Boyutlu Ki – Kare Testi (Ki-Kare İyi Uyum Testi)

Örnek: Bir hastane yıl boyunca her ay gerçekleşen poliklinik sayısının eşit olduğunu düşünmektedir (H_0). İzleyen tabloya göre bu öngörü doğru mudur?

	Ocak	Şubat	Mart	Nisan	Mayıs	Haziran	Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık	Toplam
Beklenen Poliklinik	115	115	115	115	115	115	115	115	115	115	115	115	1380
Gerçekleşen Poliklinik	110	120	80	90	105	95	120	150	120	125	145	120	1380

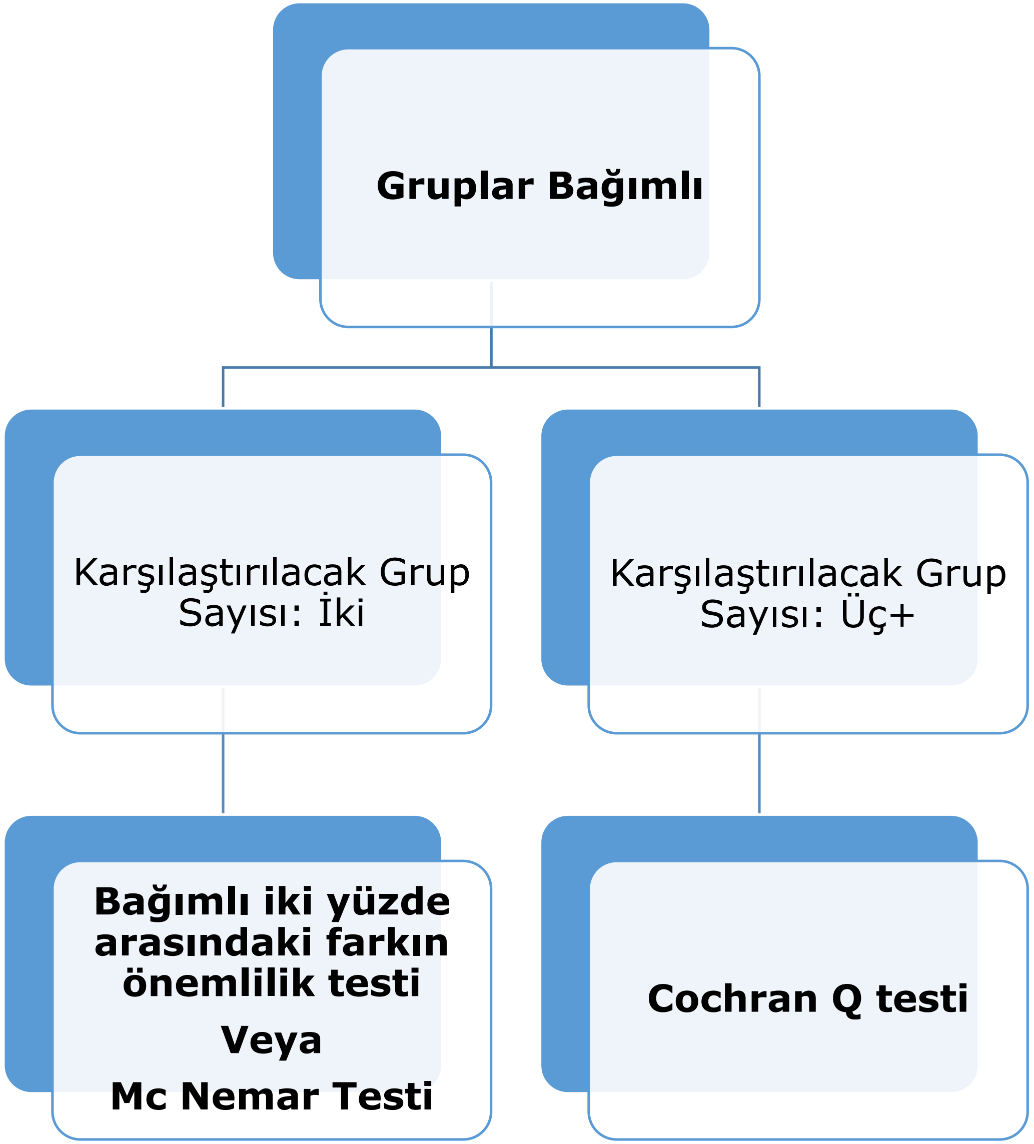
$$\chi^2_{(H)} = 40.89$$

$$\alpha = 0.05 \ \text{için} \quad \chi^2_{11,0.05} = 19.675$$

$$\chi^2_{(H)} > \chi^2_{(m-1)} \rightarrow H_0 \text{ red}$$

hastanede yıl boyunca her ay gerçekleşen poliklinik sayısının eşit değildir

Değişken Tipi: Kategorik (Sayımla Elde Edilmiş)



McNemar Testi (Bağımlı Örneklerde Ki –Kare)

- Birimlerin değişim öncesi ve sonrasında nasıl etkilendiklerini göstermek amacı ile uygulanır.
- Önce (+) olguya sahip iken sonra (-) olguya sahip olan birimlerin ya da bunun tersinde meydana gelen değişimin istatistiksel olarak anlamlı bir farklılık olup olmadığını göstermek amacı ile bu test kullanılır.

		2. Durum		
		DOĞRU	YANLIŞ	Toplam
1. Durum	DOĞRU	a	b	a+b
	YANLIŞ	c	d	c+d

Hipotezler:

H_0: İlgilenen özellik, 1. ve 2. durumda aynıdır

H_1: İlgilenen özellik, 1. ve 2. durumda değişmiştir.

$$\chi^2_{(H)} > \chi^2_{(Sd=1;\alpha=0,05)} \quad \rightarrow H_0 \text{ red.}$$

Örnek: Aşağıdaki tabloya göre bir yemek firmasınca üretilen yemeğin yapılış tarifinin değişmesi sonucu, aynı müşterilerce beğenilme durumunun değişip değişmediğini $\alpha = 0.05$ için test ediniz.

		2. Tarif		
		Beğendim	Beğenmedim	Toplam
1. Tarif	Beğendim	45	15	60
	Beğenmedim	38	2	40

Hipotezler:

H_0: Beğenme Durumu tarif ile değişmez

H_1: Beğenme Durumu tarif ile değişir

.

$$\chi^2_{(H)} = 9.98 > \chi^2_{(Sd=1;\alpha=0,05)} = 3.841$$ → H_0 red

Kolmogorov-Smirnov Testi

- Kolmogorov-Smirnov (K-S) sınaması parametrik olmayan istatistik olup iki değişik problem için hipotez sınaması olarak kullanılır:

1. *Tek örneklem K-S sınaması*: Hedef, verilmiş olan bir deneysel olasılık dağılımı gösteren örneklem verilerinin, dağılım parametreleri tam olarak bilinen tam tanımlanmış bir teorik anakitle olasılık dağılımına uyum gösterip göstermediğini sınamak. Bu tip problemde *sıfır hipotez Ho* örneklem verilerin deneysel dağılımının tam tanımlanmış bir anakitle olasılık dağılımından gelmiş olduğudur.
2. *İki örneklem K-S sınaması*: Hedef, verilmis iki tane değişik deneysel olasılık dağılımı gösteren iki orneklem veri serisinin aynı tek bir teorik anakitle olasılık dağılımından gelip gelmediğini sınamak. Bu tip problemde *sıfır hipotez Ho* ise iki orneklem verilerin deneysel dağılımlarının tek bir anakitle olasılık dağılımindan gelmiş olduğudur.

H_0: Değişken belirlenen dağılıma uyar

$K_{(H)} > K_{(H)}$ → H_0 red